Measuring psychological distress in older Aboriginal and Torres Strait Islanders Australians: a comparison of the K-10 and K-5

Bridgette J. McNamara,[1] Emily Banks,[2,3] Lina Gubhaju,[1] Anna Williamson,[2] Grace Joshy,[3] Beverley Raphael,[4] Sandra J. Eades[1]

Abstract

Objectives: To assess the cross-cultural validity of two Kessler psychological distress scales (K-10 and K-5) by examining their measurement properties among older Aboriginal and Torres Strait Islanders and comparing them to those in non-Aboriginal individuals from NSW Australia.

Methods: Self-reported questionnaire data from the 45 and Up Study for 1,631 Aboriginal and 231,774 non-Aboriginal people were used to examine the factor structure, convergent validity, internal consistency and levels of missing data of K-10 and K-5.

Results: We found excellent agreement in classification of distress of Aboriginal participants by K-10 and K-5 (weighted kappa=0.87), high internal consistency (Cronbach's alpha K-10: 0.93, K-5: 0.88), and factor structures consistent with those for the total Australian population. Convergent validity was evidenced by a strong graded relationship between the level of distress and the odds of: problems with daily activities due to emotional problems; current treatment for depression or anxiety; and poor quality of life.

Conclusions and implications: K-10 and K-5 scales are promising tools for measuring psychological distress among Aboriginal and Torres Strait Islanders aged 45 and over in research and clinical settings.

Key words: psychological distress, Aboriginal Australians, cross-cultural validity, Kesser scales, psychometric

Aboriginal* Australians (3% of the total Australian population), have significantly poorer mental health than non-Aboriginal Australians, with higher rates of hospitalisation,[1] intentional self-harm[1] and associated mortality.[2] Measurement tools that are valid and reliable for this group and allow comparison with data from other populations are vital for understanding the scope and determinants of these disparities, and for informing the provision and evaluation of mental health services.[3,4] Differences in the conceptualisation of mental health between Aboriginal Australians and other Australians,[5,6] and the language used to describe emotional states between these groups, may have implications for the validity of measurement tools; cross-cultural validity should not be automatically assumed.[4,7] Aboriginal concepts of 'social and emotional wellbeing' tend to be broader than the Western concept of 'mental health', recognising the "importance of connection to land, culture, spirituality, ancestry, family and community and how these affect the individual"[5] rather than emphasising the presence or absence of mental health conditions. Terms such as 'depression' or 'depressed' are less commonly used among some Aboriginal communities than they are in other Australian communities.[6]

The K-10 has been used in several population health surveys with Aboriginal Australians[8,9] and is routinely used in Australian clinical practice. Despite this, to the best of our knowledge, the only published data on the measurement properties of Kessler scales for Aboriginal Australians is a demonstration of the convergent validity of a modified K-5 with various mental health indicators.[1] However, the K-5 version has had limited use and is not used routinely in clinical practice or for health surveys of non-Aboriginal populations. Thus there is a considerable knowledge gap about the validity of the most commonly used scale – the K-10 – for Aboriginal Australians, which needs to be addressed.

The 45 and Up Study of New South Wales (NSW), a longitudinal cohort of people aged 45 years and over from NSW, Australia,[10] has data on a large sample of middle-aged and older Aboriginal people from a diverse range of communities across the state. Our study aimed to assess various psychometric properties of the K-10 and K-5 scales in Aboriginal participants from the 45 and Up Study and to make comparisons with non-Aboriginal individuals. Factor structure, convergent validity with other mental health indicators, internal consistency reliability of the scales and the amount of missing data were examined. We hypothesised that the properties of the K-10 for Aboriginal

* The term Aboriginal has been used to describe Aboriginal and/or Torres Strait Islander individuals, in keeping with advice from the Aboriginal Health and Medical Research Council of NSW but with respect and acknowledgement of Torres Strait Islanders as a distinct cultural group.

1. Preventative Health, Baker IDI Heart and Diabetes Institute, Victoria
2. The Sax Institute, New South Wales
3. National Centre for Epidemiology and Population Health, The Australian National University, Australian Capital Territory
4. Psychological and Addiction Medicine, The Australian National University, Australian Capital Territory
Correspondence to: Professor Sandra J. Eades, The Sax Institute, PO Box K617, Haymarket, NSW 1240; e-mail: Sandra.Eades@bakeridi.edu.au
Submitted: January 2014; Revision requested: February 2014; Accepted: June 2014
The authors have stated they have no conflict of interest.
Aust NZ J Public Health. 2014; 38:567-73; doi: 10.1111/1753-6405.12271

participants would be similar to those obtained by including only questions from the K-5, and that the properties of the scales would not differ markedly for Aboriginal and non-Aboriginal participants.

Methods

The 45 and Up Study

The Sax Institute's 45 and Up Study is a large-scale longitudinal cohort study of women and men aged 45 years and older from the general population of NSW. It includes a baseline questionnaire and linkage to routinely collected health data, with planned five-yearly repeat questionnaires and the potential for more detailed sub-studies;[10] further information is available at www.saxinstitute.org.au/our-work/45-up-study/.

Ethical approval for the current study was granted by the Aboriginal Health and Medical Research Council of New South Wales (reference 775/11) and the Human Research Ethics Committee of the University of New South Wales (reference 10186).

Participant recruitment

Individuals were randomly selected from the Medicare Australia (national health insurance) database, with over-sampling of rural areas and individuals aged 80 years and older. Residents in remote areas were completely enumerated. Participants entered the study by completing a baseline postal questionnaire and by providing written consent to have their health followed over time. The baseline questionnaire was distributed between 1 January 2006 and 31 December 2008. A total of 266,848 people were recruited to the study at the time of the October 2010 data release; of these, 266,611 had a valid age and date of entry recorded.

Aboriginal and Torres Strait Islander status was through self-identification on the baseline questionnaire. Of the total participants, 1,939 people identified as being of Aboriginal or Torres Strait Islander origin.

Data collection

The baseline postal questionnaire was used to collect socio-demographic, health and lifestyle information from participants. The baseline characteristics of Aboriginal participants have been reported previously.[11] The current study used participant data relating to: Aboriginal status; age; sex; responses to the K-10 scale for psychological distress and the K-5 items alone; questions about emotional impact and self-rated quality of life; and questions about recent treatment for anxiety and depression.

The K-10 psychological distress scale has 10 related items asking participants about the signs and symptoms of distress: "During the past 4 weeks, how often did you feel: tired out for no good reason? (Item 1); Nervous?* (Item 2); So nervous that nothing could calm you down? (Item 3); Hopeless?* (Item 4); Restless or fidgety?* (Item 5); So restless you could not sit still? (Item 6); Depressed? (Item 7); That everything was an effort?* (Item 8); So sad that nothing could cheer you up?* (Item 9); Worthless? (Item 10)". Five response options were provided for each question: "none of the time"; "a little of the time"; "some of the time"; "most of the time"; or "all of the time"; with a corresponding score of 1–5.

The current study used these items to measure psychological distress in two ways: 1) using the standard K-10 scale; and 2) using only items in the K-5 version of the scale (asterisked items 2, 4, 5, 8, 9). The modified K-5 version of the Kessler scale was developed for use in the social and emotional wellbeing module of the National Aboriginal and Torres Strait Islander Health Survey (NATSIHS) through consultation with Aboriginal people. Modifications included omission of the K-6 item "feel worthless" to avoid the potential for offence and slight word changes to other items to improve the understanding of these items ("restless or fidgety" was changed to "restless and jumpy", "hopeless" to "without hope" and "past four weeks" to "last four weeks").[1] The K-5 in our study was embedded in the K-10, and thus used the standard wording without modification.

Kessler scores can only be calculated (by summation of item scores) where there are responses for all 10 items (or five items for the K-5). Logical backfilling and imputation were used to reduce the amount of missing data for this scale.[11,12] Briefly, for item pairings 2 and 3, 5 and 6, or 8 and 9, where the latter item pair was missing but where a response of "none of the time" was recorded for the corresponding less severe item (2, 5 or 8), this response was recorded for the respective missing item. For participants with a total of nine complete items, the mean of the participant's other responses was imputed for the missing item.

Distress scores were categorised according to systems used by the Australian Bureau of Statistics (ABS) for the K-10 [Low: 10–15.99; Moderate 16–21.99; High: 22–29.99; Very High: 30–50. K-5];[13] and the NATSIHS for the K-5 [Low: 5–7.99; Moderate 8–11.99; High: 12–14.99; Very High: 15– 25].[1] Established classifications were used for each scale, despite the small difference in the lower cut-point of 'High distress', in order that the analysis would have greatest relevance for other researchers and clinicians.

The *impact of emotional problems* on daily life was measured using a series of three questions and yes/no check boxes: "During the past four weeks, have you had any of the following problems with your work or daily activities because of any emotional problems (such as being depressed or anxious)? 1) Cut down on the amount of time you spent on work or other activities 2); achieved less than you would have liked to 3); did work or other activities less carefully than usual".

Information on *recent treatment for anxiety or depression* was obtained using a question that asked about recent treatment for a number of conditions including depression and anxiety: "In the last month have you been treated for: depression or anxiety"; with a check box for "yes" for each of the conditions.

Self-rated quality of life was measured using the question: "In general, how would you rate your: Quality of life?"; with response options of excellent, very good, good, fair and poor. This variable was dichotomised into 'Poor' and 'Fair and better' for the following analyses.

Statistical analyses

Weighted Kappa was used to assess agreement in categorical distress levels. A paired t-test was used to test for a difference in the continuous distress scores by each scale. Scores were rescaled as follows to a 0–100 scale to allow comparison: K10_rescaled=(K10-10)/(50-10) * 100; K5_rescaled=(K5-5)/(25-5) * 100.

Confirmatory factor analysis (PROC CALIS with an asymmetric distribution-free weighted least squares estimation method [METHOD=ADF]) was used to assess the factor structure of the K-10 and K-5. Using participants' data without imputation, we tested the three models that were described in a review of previously published factor analyses for K-10.[14] Model 1: all 10 items loaded to a single factor 'Psychological distress'; Model 2: all 10 items load to the single factor, but with correlated error terms for the paired items; Model 3: items load to two factors, items 2, 3, 5, 6 to an 'Anxiety' factor and items 1, 4, 7-10 to a 'Depression' factor. A fourth model assessed the structure of K-5 – Model 4: all items of the K-5 loaded to a single factor 'Psychological distress'. Several fit statistics were used to determine the most appropriate structure for these population groups; the chi-square statistic,

testing the difference between the observed and expected covariance matrices, the comparative fix index (CFI), the Tucker-Lewis Index (also known as the Non-Normed Fit index) and the root mean square error of approximation (RMSEA).[15,16] Further post-hoc analysis was done to assess measurement invariance of the K10 scale between the two groups using two nested multi-group confirmatory factor models. This analysis, conducted using Mplus 6.12 Software,[17] used polychoric correlations and robust weighted least squares estimation (WLSMV) – appropriate for ordered categorical indicators. Item 9 "feel effort" was fixed to set the scale for the latent variable (distress). Scalar invariance was tested, by comparing the fit (using DIFFTEST option and fit indices) of a model where factor loadings and thresholds for the Aboriginal group were constrained to be the same as the non-Aboriginal group, to a first model where these parameters were freely estimated.

Convergent validity was measured in this study by examining the relationship between level of distress by the K-10 and K-5 and: 1) the impact of emotional problems; 2) recent treatment for depression or anxiety; and 3) poor self-rated quality of life. The proportions of Aboriginal and non-Aboriginal participants with the above outcomes by level of psychological distress are presented, and unconditional logistic regression was used to calculate age- and sex- adjusted odds ratios (ORs) and 95% confidence intervals (CIs) for each outcome.

Internal consistency of the K-10 and K-5 was assessed by Cronbach's alpha, and by examining correlations between individual item scores and the total scores (item-total consistency). An alpha of 0.7 or greater and item-total correlations of greater than 0.5 were designated *a priori* to denote an acceptable level of internal consistency.

We examined the proportion of missing/invalid data for individual items of the K-10 and for the entire scale. Of particular interest was the proportion of missing data for the items about which concerns were raised during the development of the K-5.[1] Socio-demographic characteristics of the Aboriginal participants with missing K-10 data after backfilling and imputation (who were therefore excluded from the reliability and validity analysis) were compared to those included in the study. Sensitivity analyses were also conducted, repeating the main analyses for participants with complete data (i.e. no backfilling or imputation). All statistical analyses (except the multi-group CFA) were undertaken using SAS software version 9.3 (SAS Institute Inc, Cary, NC, USA).

Results

Complete data for all K-10 items were available for 1,589 (82%) of the Aboriginal participants and 227,063 (87%) of the non-Aboriginal participants. Following imputation, this increased to 1,631 (84%) and 231,774 (89%), respectively. Sensitivity analysis showed that results did not differ with the inclusion of imputed data so the analysis presented here, except the factor structure analysis, used all available data.

Levels of distress by K-10 and K-5

The levels of psychological distress measured by K-10 and K-5 are shown in Table 1. The percentage of participants experiencing high to very high levels of psychological distress was 20.9% (95%CI 18.9-22.9) by K-10 and 19.3% (17.3-21.1%) by K-5. The weighted Kappa statistic was 0. 87 (0.86-0.89), indicating an excellent level of agreement. Complete agreement in the level of distress was found in 85% of participants. There were no Aboriginal participants whose classification differed by more than one category. A paired t-test of the K-10 and rescaled K-5 scores found no significant difference between the distress scores for each scale among Aboriginal participants (p=0.54).

Fewer non-Aboriginal participants had high/very high levels of distress than Aboriginal participants; 7.5% (7.37-7.59) and 7.0% (6.86-7.07) of non-Aboriginal participants had high/very high distress by K-10 and K-5, respectively. Compared to Aboriginal participants, there was a greater level of exact agreement in classification (90% of cases) for non-Aboriginal participants, and a high but slightly lower level of agreement overall with a weighted Kappa of 0.82 (0.82-0.82), with four (0.001%) participants differing by more than one category. The proportion of participants with very high distress was statistically significantly greater by K-5 than by K-10, and a paired t-test showed a higher average score by K-5 (p<0.0001); however, the mean difference is not likely to be clinically significant (less than 0.5% of the total range of scores).

Confirmatory factor analysis

The most appropriate factor structure for K-10 in both Aboriginal and non-Aboriginal participants (Table 2) was all items loading to a single 'Psychological distress' factor and correlated error terms between the paired items (Model 2). All fit statistics for Model 2 were within the accepted levels for a well- to reasonably-fitted model (>0.9 CFI & TLI; less than 0.08 RMSEA),[15,16] with the exception of the chi-square statistic; however, this statistic is extremely sensitive to large sample sizes.[16]

Configural invariance (the same factor structure for both groups) of Model 2 was confirmed by the multi-group analysis (CFI 0.980, TLI 0.972, RMSEA 0.075 [90%CI 0.075-0.076]). When the factor loadings and thresholds were constrained to be equal between Aboriginal and non-Aboriginal groups, the model produced the following fit statistics (CFI 0.988, TLI 0.991, RMSEA 0.043 [90%CI 0.043-0.044]). Difference testing of the chi-square statistic using the appropriate DIFFTEST option in Mplus[17] found differences in the models to be statistically significant (chi-square 72.92, df 47, p=0.009), suggesting that scalar invariance should be rejected, and that participants might be interpreting and/or responding to the items differently. Given the large sample size of our study, it is probably more appropriate to include consideration of the size of the variation, as well as measures of statistical significance.[18] In investigation of the freely estimated factor loadings, we found that the estimates did not differ significantly between the two groups (based on 99%CIs calculated from the standard errors); however, thresholds were lower for Aboriginal participants across

Table 1: Levels of psychological distress in Aboriginal and non-Aboriginal participants using K-10 and K-5.

Level of distress	n	K-10 % (95%CI)	n	K-5 % (95%CI)
Aboriginal				
Low	971	60 (57-62)	931	57 (55-59)
Moderate	319	20 (18-21)	386	24 (22-26)
High	191	12 (10-13)	137	8 (7-10)
Very High	150	9 (8-11)	177	11 (9-12)
Non-Aboriginal				
Low	177,844	77 (77-77)	165,541	71 (71-72)
Moderate	36,593	16 (16-16)	50,046	22 (21-22)
High	12,380	5 (5-5)	9,675	4 (4-4)
Very High	4,957	2 (2-2)	6,475	3 (3-3)

all items (see Supplementary Tables 1 and 2). The lower thresholds among Aboriginal participants were consistent across the scale items and investigation of modification indices for the constrained model did not indicate that releasing any particular factor loading or threshold would significantly improve the model fit (no indices for factor loadings or thresholds reached a 0.05 level of significance).

Convergent validity

The proportion and age- and sex-adjusted odds of experiencing problems with work or daily activities because of emotional problems during the past four weeks increased with increasing levels of psychological distress in Aboriginal and non-Aboriginal participants (Table 3). Those with higher levels of distress measured by both the K-10 and K-5 were significantly more likely to have: 1) reduced time spent on work or other activities; 2) achieved less than they would have liked to; and 3) done their work or other activities less carefully than usual, when compared to participants of that group with low levels of distress. Among those with very high levels of distress, the proportion of Aboriginal participants experiencing these impacts were: 75.8%, 83.9% and 68.4% by the K-10; and 73.7%, 84.6% and 66.7% by the K-5, respectively. The odds ratios for the emotional impacts for participants with higher distress levels compared to low distress did not differ significantly between the Aboriginal and non-Aboriginal groups and ranged from 20 to 44 for the highest to the lowest distress levels.

The proportion of Aboriginal and non-Aboriginal participants who reported recent treatment for depression or anxiety was also significantly greater in participants with higher levels of distress than in those with low distress (Table 4). The odds of treatment for anxiety or depression in Aboriginal participants experiencing high and very high distress compared to those with low levels of distress were: 10.1 (95%CI 6.7-15.2) and 30.0 (19.3-46.6) by K-10; or 9.6 (6.2-14.9) and 23.0 (15.2-34.6) by K-5, respectively.

A total of 5.2% of Aboriginal participants reported poor quality of life, compared to 1.5% of non-Aboriginal participants. The proportion of participants who rated their quality of life as poor increased with increasing levels of distress (Table 5), with 35% of Aboriginal people and 24% of non-Aboriginal people experiencing very high levels of distress by K-10 rating their quality of life as poor, or 31% and 21% by K-5.

Internal consistency of the scales

The internal consistencies of the K-10 and K-5 scales were high for Aboriginal and non-Aboriginal participants: for K-10 the Cronbach's Alpha was 0.93 for Aboriginal and 0.89 for non-Aboriginal; and for K-5 Cronbach's alpha was 0.88 in Aboriginal and 0.82 in non-Aboriginal participants. Correlations between all scale items and the total scores were high, ranging from 0.57-0.83 in the Aboriginal group and 0.51-0.75 in the non-Aboriginal group for the K-10, to between 0.67 and 0.77 and 0.67 and 0.78 for the K-5. The first item in the K-10 ("During the past 4 weeks about

Table 2: Confirmatory factor analysis fit statistics for the four models tested in Aboriginal and non-Aboriginal participant groups.

Model	Chi-square	df	p	CFI	TLI	RMSEA	(RMSEA 90%CI)
Aboriginal							
1	393.11	35	<0.0001	0.96	0.94	0.082	0.074–0.089
2	243.10	32	<0.0001	0.97	0.96	0.066	0.058–0.073
3	316.23	34	<0.0001	0.96	0.95	0.074	0.066–0.081
4 (K-5)	24.21	5	0.0002	0.99	0.98	0.050	0.031–0.071
Non-Aboriginal							
1	46,573.71	35	<0.0001	0.86	0.82	0.077	0.077–0.078
2	30,618.38	32	<0.0001	0.91	0.87	0.066	0.065–0.066
3	37,905.19	34	<0.0001	0.89	0.85	0.071	0.070–0.071
4 (K-5)	5,204.64	5	<0.0001	0.94	0.88	0.068	0.067–0.070

Table 3: Proportion and adjusted odds of experiencing problems with work or daily activities because of emotional problems during the past four weeks by level of psychological distress using K-10 and K-5 in Aboriginal and non-Aboriginal participants.

Distress level	K-10		K-5	
	% (n)	OR adjusted age and sex	% (n)	OR adjusted age and sex
Aboriginal				
Reduced time you spent on work or other activities				
Low	6.6 (63)	1.0	6.3 (57)	1.0
Moderate	30.5 (93)	6.0 (4.2-8.6)	30.2 (112)	6.3 (4.5-9.0)
High	58.1 (104)	19.1 (12.9-28.4)	58.9 (76)	21.0 (13.5-32.8)
Very High	75.8 (100)	42.0 (26.1-67.6)	73.7 (115)	40.2 (25.6-62.9)
Achieved less that you would have liked to				
Low	20.0 (191)	1.0	20.0 (183)	1.0
Moderate	62.4 (194)	6.7 (5.0-8.9)	58.7 (222)	5.8 (4.4-7.5)
High	80.9 (148)	17.3 (11.5-25.9)	80.3 (106)	22.2 (14.0-35.2)
Very High	83.9 (115)	21.2 (13.0-34.5)	84.6 (137)	22.3 (14.0-35.2)
Did work or other activities less carefully than usual				
Low	7.1 (67)	1.0	7.1 (64)	1.0
Moderate	30.9 (95)	5.9 (4.2-8.3)	31.0 (115)	6.0 (4.2-8.4)
High	57.3 (102)	17.9 (12.1-26.5)	55.1 (70)	16.3 (10.5-25.2)
Very High	68.4 (91)	28.1 (18.0-44.0)	66.7 (106)	26.3 (17.2-40.1)
Non-Aboriginal				
Reduced time you spent on work or other activities				
Low	3.8 (6616)	1.0	3.5 (5755)	1.0
Moderate	24.3 (8601)	8.1 (7.8-8.3)	20.8 (10132)	7.1 (6.9-7.4)
High	49.8 (5903)	25.0 (23.9-26.1)	47.4 (4411)	24.5 (23.3-25.7)
Very High	63.6 (2968)	44.1 (41.3-47.1)	61.8 (3786)	43.9 (41.5-46.6)
Achieved less that you would have liked to				
Low	17.4 (30587)	1.0	16.4 (26719)	1.0
Moderate	55.2 (19833)	6.0 (5.8-6.1)	50.2 (24672)	5.3 (5.2-5.4)
High	77.3 (9327)	16.6 (15.9-17.4)	75.3 (7114)	16.2 (15.4-17.0)
Very High	80.2 (3796)	19.8 (18.4-21.3)	80.8 (5028)	22.4 (20.9-23.9)
Did work or other activities less carefully than usual				
Low	4.9 (8607)	1.0	4.5 (7251)	1.0
Moderate	28.0 (9929)	7.4 (7.2-7.6)	24.5 (11935)	6.8 (6.7-7.1)
High	51.6 (6135)	20.2 (19.4-21.1)	49.4 (4604)	20.6 (19.6-21.6)
Very High	62.1 (2888)	31.0 (29.1-33.0)	61.6 (3765)	33.7 (31.9-35.7)

how often did you feel: Tired out for no good reason?") had the weakest correlation with the total in both participant groups. This item is not included in the K-5.

Missing and invalid responses

The proportion of missing and invalid data for the K-10 scale items can give an indication of the acceptability of the individual items and of the total scale.

The proportion of missing/invalid data for each of the individual items was similar for all items, with the exception of the first question, which had a lower proportion of missing/invalid data in both participant groups (Table 6). The proportion of missing/invalid data was significantly greater for Aboriginal than non-Aboriginal participants for all scale items ($p<0.0001$). The proportion of missing data was also higher for Aboriginal participants across the range of other questions on the baseline questionnaire (mean 7.1% compared to 4.5%).

K-10 scores could not be calculated for 308 (16%) of the Aboriginal participants and 28,210 (11%) of non-Aboriginal participants. Analysis of the characteristics of the Aboriginal group with missing K-10 scores showed that they were significantly older ($p<0.0001$); 12.0% of the missing data group were aged over 80 years compared to 1.7% in the complete data group. They had significantly lower levels of education – 42.2% had no formal education qualifications compared to 26.2% in the group with complete data – and had poorer self-rated health ($p<0.0007$) and self-rated quality of life ($p=0.0007$). There were no significant differences in sex ($p=0.14$), recent treatment for depression or anxiety ($p=0.22$), or in SEIFA area level of disadvantage between the missing data or complete data groups ($p=0.26$). The proportion of missing data on other questions was also higher for the group missing K-10 data than those with complete K-10 data (3.5 times higher, 17% vs. 5%).

Using the K-5 questions alone there were 29 more participants (1.5% of total Aboriginal participants) who had complete data for this scale. These participants were not utilised in the analysis of validity or reliability, to allow comparison to the K-10. Completing enough scale items to calculate the K-5 but not the K-10 was associated with higher levels of distress ($p<0.001$), but not with statistically significant differences in age ($p=0.06$) or sex ($p=0.80$).

Discussion

Kessler distress scales have been used increasingly in population health surveys and clinical practice; however, the properties of these tools in groups such as the Australian Aboriginal population are only beginning to be examined. To our knowledge, this is the first study to examine measurement properties of the K-10 for the Australian Aboriginal population.

This study supports the use of both the K-10 and the K-5 for middle-aged and older Aboriginal people, and provides evidence on aspects of the validity and reliability of the scales for this population. We found that: 1) levels of distress did not differ materially when measured by K-10 or by K-5; 2) the factor structure of the K-10 was similar to that reported for the general Australian population and there was no clear indication of measurement differences; 3) the level of distress showed a strong relationship with the impacts of emotional problems, with current treatment for depression or anxiety and with poor self-rated quality of life when measured by either scale; 4) both scales had high internal consistency; and 5) no individual item had a higher proportion of missing data, but that Aboriginal participants had a higher proportion of missing data for all scale items and for other questions on the baseline questionnaire.

The unweighted proportion of Aboriginal participants aged 45 and older experiencing high to very high psychological distress by K-10 in this study was 20.9% (95%CI 18.9-22.9), which was comparable to that reported in the NSW population health survey and the National Aboriginal and Torres Strait Islander Health Survey. The NSW population health survey (2002–2005) found that 28.0% (95%CI 19.6-36.5) of Aboriginal participants aged 45–54 years; 21.3% (13.0-29.7) of those aged 55-64; and 9.3% (3.6-14.9) of those aged 65 and over had high to very high levels of distress, as measured by the K-10.[8,9] The NATSIHS (2004–2005) found 29.4% (95%CI 24.4-34.5) of Aboriginal and Torres Strait Islander people aged 45–54 had high to very high levels of distress on the K-5 scale, and 23% (18.4-27.5) of those aged 55 and over.[1]

Results of the confirmatory factor analysis of the K-10 and K-5 for both the Aboriginal and non-Aboriginal group revealed a similar structure to that previously reported for the Australian population. All the goodness-of-fit statistics for each group supported the single

Table 4: Proportion and adjusted odds of treatment for anxiety or depression during the past month by level of psychological distress using K-10 and K-5 in Aboriginal and non-Aboriginal participants.

	K-5		K-5	
Psychological distress	% (n)	OR for treatment for anxiety or depression-adjusted age and sex	% (n)	OR for treatment for anxiety or depression-adjusted age and sex
Aboriginal				
Low	5.9 (57)	1.0	6.4 (59)	1.0
Moderate	19.4 (62)	4.0 (2.7-5.9)	17.9 (69)	3.4 (2.3-4.9)
High	36.7 (70)	10.1 (6.7-15.2)	38.0 (52)	9.6 (6.2-14.9)
Very High	64.0 (96)	30.0 (19.3-46.6)	59.3 (105)	23.0 (15.2-34.6)
Non-Aboriginal				
Low	3.8 (6532)	1.0	3.6 (5973)	1.0
Moderate	15.4 (5639)	4.7 (4.5-4.8)	13.5 (6761)	4.0 (3.9-4.2)
High	32.7 (4043)	12.4 (11.9-13.0)	30.7 (2970)	11.5 (11.0-12.1)
Very High	50.4 (2499)	26.2 (24.6-27.9)	46.4 (3007)	22.7 (21.4-23.9)

Table 5: Proportion and adjusted odds of poor quality of life by level of psychological distress using K-10 and K-5 in Aboriginal and non-Aboriginal participants.

	K-10		K-5	
Psychological distress	% (n)	OR of poor quality of life-adjusted age and sex	% (n)	OR of poor quality of life-adjusted age and sex
Aboriginal				
Low	0.9 (8)	1.0	1.1 (9)	1.0
Moderate	3.0 (9)	3.7 (1.4-9.7)	2.0 (7)	2.0 (0.7-5.3)
High	8.4 (18)	10.6 (4.4-25.6)	9.9 (13)	10.9 (4.5-26.2)
Very High	34.6 (49)	71.0 (31.7-159.1)	31.3 (50)	49.7 (23.3-106.2)
Non-Aboriginal				
Low	0.3 (470)	1.0	0.3 (450)	1.0
Moderate	2.0 (715)	8.3 (7.4-9.3)	1.8 (839)	7.0 (6.2-7.9)
High	7.8 (912)	35.1 (31.3-39.4)	6.8 (625)	30.0 (26.5-34.0)
Very High	23.9 (1120)	137.1 (122.2-153.9)	21.3 (1303)	114.9 (102.6-128.6)

Table 6: Proportion (number) of missing/invalid responses for K-10 in Aboriginal and non-Aboriginal participants. Questions included in K-5 are bolded.

K-10 Question	Aboriginal	Non-Aboriginal	*p*-value
1. Tired out for no good reason	7% (128)	4% (11200)	<0.0001
2. Nervous	13% (257)	9% (23690)	<0.0001
3. So nervous that nothing could calm you down?	14% (277)	10% (26494)	<0.0001
4. Hopeless	13% (256)	10% (25615)	<0.0001
5. Restless or fidgety	13% (248)	9% (24039)	<0.0001
6. So restless that you could not sit still	13% (252)	10% (26241)	<0.0001
7. Depressed	12% (225)	9% (22430)	<0.0001
8. So sad that nothing could cheer you up[a]	13% (257)	10% (24978)	<0.0001
9. Everything was an effort	12% (235)	9% (22210)	<0.0001
10. Worthless	13% (253)	10% (25169)	<0.0001

a. In Version 1, 'too depressed' was used instead of 'so sad', so for version 1 questionnaires, (n=37,000 in total, n=274 Aboriginal and Torres Strait Islander).

factor structure with correlated error terms found by Sunderland et al (2012) for their community sample (2007 Australian National Survey of Mental Health and Wellbeing) for the K-10[14] and a simple single factor structure of the K-5 as described previously for the K-6.[14] Investigation of measurement invariance of the K-10 did not indicate specific differences in the way the indicator items were measuring distress between the groups. We found similar factor loadings but lower thresholds for all K-10 scale items, suggesting that while Aboriginal participants were interpreting the items in the same way, they appeared to be endorsing all scale item categories at a slightly lower level of 'distress' than the non-Aboriginal participants. This differential endorsement may be contributing in a small way to the higher K-10 distress scores of the Aboriginal group. Future studies that examine the clinical predictive validity of the K-10 in both population groups are needed to provide further insight into the clinical significance of this finding.

The results of this study demonstrate equivalent convergent validity of the K-10 and K-5 psychological distress scales with the relevant mental health variables examined. When assessing the convergent validity of the scales, we considered the two primary purposes for which the Kessler scales were developed: first, to assess the severity of psychological distress at a population level; and, second, to be a screening tool to detect likely cases of mental health disorders (particularly depression and anxiety) in population samples.[19] We found a strong graduated relationship between the level of distress measured by the Kessler scales and the proportion of Aboriginal and non-Aboriginal participants reporting impacts of distress on their daily lives, and a poor subjective rating of quality of life.

The proportions of participants reporting recent treatment for depression or anxiety increased with the level of distress. The fact that 64% or 59% of Aboriginal participants with very high distress measured by K-10 or K-5 reported receiving treatment for depression or anxiety over the last month is encouraging. While barriers to accessing mental health treatment have been reported for Aboriginal people,[20,21] we found the proportion of Aboriginal participants in this study cohort who reported recent treatment was greater than non-Aboriginal participants across the spectrum of distress; only 50% or 46% of non-Aboriginal participants with very high distress by K-10 or K-5 reported recent treatment. Greater health-service utilisation in a more highly educated study population compared to the general Aboriginal population[11] may be contributing higher proportions of reported treatment.

Examining the internal consistency of each of the scales, we found no difference in the consistency of either scale – both had a very high consistency. Item 1 of the K-10 "feel tired" showed the weakest correlation with the total scale scores. This item also had the lowest proportion of missing data. This was the case for both Aboriginal and non-Aboriginal participants. Furthermore, the amount of missing data did not differ between each scale item for Aboriginal participants (with the exception of the first item with a lower proportion), suggesting that no particular scale item appeared to be problematic to the extent that participants missed the individual item, including the item "feel worthless" that was omitted from the K-5 during development due to concerns about its acceptability.

Aboriginal participants had greater proportions of missing data for each Kessler scale item, and across the entire postal questionnaire. The proportion for the Kessler scale was 1.5 times that of non-Aboriginal participants and the proportion for the range of other questionnaire items was 1.6 times, suggesting that the higher levels of missing data among Aboriginal participants represent a general pattern and is not specific to the Kessler scale. The exact reason for the greater amount of missing data for Aboriginal participants is not known. It stands to reason that those who are experiencing high levels of distress may also be more likely to miss Kessler scale items and other questions. The fact that the 1.5% of Aboriginal participants who completed sufficient items to calculate a K-5 distress score, but not sufficient to calculate the K-10 score, had significantly higher levels of distress than those with complete K-10 data supports this reasoning, and must thus be considered a limitation of the current study. Third party data collection would be less likely to have this limitation. However, collecting the information in person or on the phone may result in other biases. It has been suggested that social desirability bias – participants not wanting to express emotional 'weakness' to a non-medical interviewer – may lower the estimates of distress when third party collection is used.[22]

Measurement properties of the K-6 were also examined by the authors as a post-hoc analysis. The K-6 is a commonly used six-item version of Kessler scale that includes items of the K-5 as well as the item "feel worthless". This scale demonstrated very similar validity and reliability to the K-5 (results provided upon request).

Study limitations

The collection of data using postal questionnaires enabled the inclusion of middle-aged and older Aboriginal people from diverse communities across NSW, but does have limitations; a response rate of 18% (among all questionnaire recipients of the 45 and Up Study) is typical of cohort studies of this nature. Also typical is the 'healthy cohort effect' (more health conscious and educated and less distressed individuals tend to be more likely to join such studies). Nevertheless, provided appropriate measures are used and sufficient heterogeneity of exposure is present, such studies provide robust estimates based on internal comparisons, including the measures of exposure-outcome relationships [23] and methodological measures examined here. The representativeness of the entire sample population is vital for prevalence estimates. For studies of the type presented here, it is considered less important than the demonstrated internal validity of the study in terms of measuring association.[24-27]

There is also the potential for misclassification of Aboriginal status if the question was misread or misinterpreted by participants, or due to routine data entry errors. Such errors would affect estimates for the Aboriginal sub-group to a greater degree, due to the small numbers in this group. On the other hand, Aboriginal and Torres Strait Islander participants may be more likely to self-identify on a confidential questionnaire than in other situations where there may be fear of discriminations or differential treatment, so it is possible we have a more inclusive sample than other some other health datasets.

As information on current treatment was based on self-report alone, the accuracy of this variable has not been verified and is the subject of ongoing research. Any misclassification in this area would tend, in general, to lead to a dilution of observed effects.

The 45 and Up Study questionnaire contained the standard K-10 scale, without the word changes of the NATSIHS K-5 version. The comparison of the scales undertaken here includes the assumption that asking the entire 10 items of the K-10 in the standard form has not influenced the responses of the questions contained in the K-5.

As the 45 and Up Study questionnaire was in English, the results of this study may only be applicable to the English-speaking proportion of the Aboriginal and non-Aboriginal populations. Similarly, the study sample was restricted to people aged 45 years and over; while we are not aware of any studies that would suggest differences in the validity of the scales for younger Aboriginal Australians compared to those over the age of 45, further study is needed to assess the validity of the Kessler scales in Aboriginal adolescents and young adults.

Implications

The establishment of robust tools for measuring mental health that are culturally appropriate but also allow comparability with other populations is essential to understand the disparities in mental health between Aboriginal and non-Aboriginal Australians. While some aspects of the reliability and validity of the K-10 and K-5 are yet to be explored (e.g. clinical predictive validity, test-retest reliability, and ability to detect change over time), the results of the current study suggest both scales are promising tools for measuring psychological distress among middle-aged and older Aboriginal people. The choice of tool would depend on the research or clinical setting, and the need for comparability or for brevity in the scale.

Acknowledgements

We acknowledge the support of the NSW Aboriginal Health and Medical Research Council (AHMRC) and the AHMRC ethics committee, and advice provided by Professor Cate D'Este regarding the multi-group confirmatory factor analysis.

This research was completed using data collected through the 45 and Up Study (www.saxinstitute.org.au), managed by the Sax Institute in collaboration with major partner Cancer Council NSW; and partners: the National Heart Foundation of Australia (NSW Division); NSW Ministry of Health; beyondblue; Ageing, Disability and Home Care, Department of Family and Community Services; the Australian Red Cross Blood Service; and UnitingCare Ageing. We thank the many thousands of Aboriginal and non-Aboriginal people participating in the 45 and Up Study.

Bridgette McNamara is supported by a National Health and Medical Research Council Early Career Fellowship. Baker IDI Heart and Diabetes Institute is supported in part the Victorian Government OIS Program. Sandra Eades and Emily Banks are supported by the National Health and Medical Research Council of Australia. Anna Williamson is supported by a National Health and Medical Research Council of Australia Training Fellowship.

References

1. Australian Institute of Health and Welfare. *Measuring the Social and Emotional Wellbeing of Aboriginal and Torres Strait Islander Peoples*. Catalogue No.: IHW 24. Canberra (AUST): AIHW; 2009.
2. Vos T, Baker B, Stanley L, Lopez A. *The Burden of Disease and Injury in Aboriginal and Torres Strait Islander Peoples 2003*. Brisbane (AUST): University of Queensland School of Population Health; 2007.
3. National Aboriginal and Torres Strait Islander Health Council. *National Strategic Framework for Aboriginal and Torres Strait Islander Health 2003-2013: Framework for Action by Governments*. Canberra (AUST): NATSIHC; 2004.
4. Beals J, Manson SM, Mitchell CM, Spicer P, Team A-S. Cultural specificity and comparison in psychiatric epidemiology: Walking the tightrope in American Indian research. *Cult Med Psychiatry*. 2003;27(3):259-89.
5. Zubrick S, Dudgeon P, Gee G, et al. Social Determinants of Aboriginal and Torres Strait Islander Social and Emotional Wellbeing. In: Purdie N, Dudgeon P, Walker R, editors. *Working Together: Aboriginal and Torres Strait Islander Mental Health and Wellbeing Principles and Practice*. Canberra (AUST): Commonwealth Department of Health and Ageing; 2010.
6. Brown A, Scales U, Beever W, Rickards B, Rowley K, O'Dea K. Exploring the expression of depression and distress in aboriginal men in central Australia: A qualitative study. *BMC Psychiatry*. 2012;12:97.
7. Williamson A, Andersen M, Redman S, et al. Measuring mental health in Indigenous young people: A review of the literature from 1998-2008. *Clin Child Psychol Psychiatry*. 2014;19(2):260-72.
8. Centre for Epidemiology and Research. *2002-2005 Report on Adult Aboriginal Health from the New South Wales Population Health Survey*. Sydney (AUST): New South Wales Department of Health; 2006.
9. Jorm AF, Bourchier SJ, Cvetkovski S, Stewart G. Mental health of Indigenous Australians: A review of findings from community surveys. *Med J Aust*. 2012;196:118-21.
10. Banks E, Redman S, Jorm L, et al. Cohort profile: the 45 and up study. *Int J Epidemiol*. 2008;37(5):941-7.
11. Gubhaju L, McNamara BJ, Banks E, et al. The overall health and risk factor profile of Australian Aboriginal and Torres Strait Islander participants from the 45 and up study. *BMC Public Health*. 2013;13(1):661.
12. Byles JE, Gallienne L, Blyth FM, Banks E. Relationship of age and gender to the prevalence and correlates of psychological distress in later life. *Int Psychogeriatr*. 2012;24(6):1009-18.
13. Australian Bureau of Statistics. *Information Paper: Use of Kessler Psychological Distress Scale in ABS Health Surveys, Australia, 2001*. Canberra (AUST): ABS; 2003.
14. Sunderland M, Mahoney A, Andrews G. Investigating the factor structure of the Kessler Psychological Distress Scale in community and clinical samples of the Australian population. *J Psychopathol Behav Assess*. 2012;34(2):253-9.
15. Hu-Li J, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6(1):1-55.
16. Fan X, Thompson B, Wang L. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Struct Equ Modeling*. 1999;6(1):56-83.
17. Muthén LK, Muthén BO. *Mplus Users' Guide*. 6th ed. Los Angeles (CA): Muthén and Muthén; 2010.
18. Brown TA. Chapter 9: Data Issues in CFA: Missing, Non-normal and Categorical Data. In: Brown TA, editor. *Confirmatory Factor Analysis for Applied Research*. New York (NY): The Guilford Press; 2006. p. 378-410.
19. Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med*. 2002;32(6):959-76.
20. Westerman T. Engagement of Indigenous clients in mental health services: What role do cultural differences play? *Aust e-J Adv Ment Health*. 2004;3(3):1-6.
21. Brown A, Blashki G. Indigenous male health disadvantage--linking the heart and mind. *Aust Fam Physician*. 2005;34(10):813-19.
22. Wooden M. *Use of the Kessler Psychological Distress Scale in the HILDA Survey*. HILDA Project Discussion Paper Series No.: 2/09. Melbourne (AUST): University of Melbourne; 2009.
23. Mealing NM, Banks E, Jorm LR, Steel DG, Clements MS, Rogers KD. Investigation of relative risk estimates from studies of the same population with contrasting response rates and designs. *BMC Med Res Methodol*. 2010;10:26.
24. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42(4):1012-4.
25. Elwood JM. Commentary: On representativeness. *Int J Epidemiol*. 2013;42(4):1014-15.
26. Richiardi L, Pizzi C, Pearce N. Commentary: Representativeness is usually not necessary and often should be avoided. *Int J Epidemiol*. 2013;42(4):1018-22.
27. Nohr EA, Olsen J. Commentary: Epidemiologists have debated representativeness for more than 40 years-has the time come to move on? *Int J Epidemiol*. 2013;42(4):1016-17.

Supporting Information

Additional supporting information may be found in the online version of this article:

Supplementary Table 1: Freely-estimated factor loadings from multi-group CFA (99%CIs).

Supplementary Table 2: Freely-estimated thresholds and standard errors from multi-group CFA. Model fixed mean of 'distress' to zero, variance of 'distress' was not fixed.